

BILINEAR FORMS WITH KLOOSTERMAN AND GAUSS SUMS

IGOR E. SHPARLINSKI

ABSTRACT. We obtain several estimates for bilinear form with Kloosterman sums. Such results can be interpreted as a measure of cancellations amongst with parameters from short intervals. In particular, for certain ranges of parameters we improve some recent results of Blomer, Fouvry, Kowalski, Michel and Milićević and also of Fouvry, Kowalski and Michel. In particular, we improve the bound on the error term in the asymptotic formula for mixed moments of L -series associated with Hecke eigenforms.

1. INTRODUCTION

1.1. Background and motivation. Let q be a positive integer. We denote the residue ring modulo q by \mathbb{Z}_q and denote the group of units of \mathbb{Z}_q by \mathbb{Z}_q^* .

For integers m and n we define the Kloosterman sum

$$\mathcal{K}_q(m, n) = \sum_{x \in \mathbb{Z}_q^*} \mathbf{e}_q(mx + n\bar{x}),$$

where \bar{x} is the multiplicative inverse of x modulo q and

$$\mathbf{e}_q(z) = \exp(2\pi iz/q).$$

Given a set $\mathcal{M} \subseteq \mathbb{Z}_q^*$, an interval $\mathcal{J} = \{L+1, \dots, L+N\} \subseteq [1, q-1]$ of N consecutive integers and a sequence of weights $\mathcal{A} = \{\alpha_m\}_{m \in \mathcal{M}}$, we define the weighted sums of Kloosterman sums

$$\mathcal{S}_q(\mathcal{A}; \mathcal{M}, \mathcal{J}) = \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{J}} \alpha_m \mathcal{K}_q(m, n).$$

By the Weil bound we have

$$|\mathcal{K}_q(m, n)| \leq q^{1/2+o(1)},$$

see [14, Corollary 11.12]. Hence we immediately obtain

$$(1.1) \quad |\mathcal{S}_q(\mathcal{A}; \mathcal{M}, \mathcal{J})| \leq Nq^{1/2+o(1)} \sum_{m \in \mathcal{M}} |\alpha_m|.$$

2010 *Mathematics Subject Classification.* 11D79, 11L07.

Key words and phrases. Kloosterman sums, cancellation, bilinear form.

We are interested in studying cancellations amongst Kloosterman sums and thus in improvements of the trivial bound (1.1).

We remark that if $q = p$ is prime, then making the change of variable $x \mapsto nx \pmod{p}$, one immediately observes that $\mathcal{K}_p(mn, 1) = \mathcal{K}_p(m, n)$, thus we also have

$$\mathcal{S}_p(\mathcal{A}; \mathcal{M}, \mathcal{J}) = \sum_{m \in \mathcal{M}} \sum_{n \in \mathcal{J}} \alpha_m \mathcal{K}_p(mn, 1).$$

If furthermore $\mathcal{M} = \mathcal{I} = \{K+1, \dots, K+M\} \subseteq [1, p-1]$ is an interval of M consecutive integers, we obtain the sums $\mathcal{S}_p(\mathcal{A}; \mathcal{I}, \mathcal{J})$, which have been studied in recent works of Blomer, Fouvry, Kowalski, Michel and Milićević [1], Fouvry, Kowalski and Michel [8] and Shparlinski and Zhang [24].

We have to stress that the most important part of the very deep work of Blomer, Fouvry, Kowalski, Michel and Milićević [1] is establishing the connection between bilinear forms with Kloosterman sums and mixed moments of L -series associated with Hecke eigenforms. However improving one of their ingredients allows us to improve one of their main results.

The sums $\mathcal{S}_p(\mathcal{A}; \mathcal{I}, \mathcal{J})$, including the special case of sums without weights, that is, $\mathcal{S}_p(\{1\}_{m=1}^M; \mathcal{I}, \mathcal{J})$, as well as some other related sums, such as

$$\sum_{m \in \mathcal{I}} \sum_{n \in \mathcal{J}} \alpha_m \beta_n \mathcal{K}_p(mn, 1) \quad \text{and} \quad \sum_{m \in \mathcal{I}} \mathcal{K}_p(m, 1),$$

(where $\{\beta_n\}_{n \in \mathcal{J}}$ is another sequence of weights) appear in some applications and are also of independent interest, we refer to [1, 2, 8, 9, 18, 24, 26] for a wide range of various applications and further references. We also recall recent results of [19, 17, 25] when cancellations among Kloosterman sums are studied for moduli of special arithmetic structure.

Here we consider more general sums $\mathcal{S}_q(\mathcal{A}; \mathcal{M}, \mathcal{J})$ extending the previously studied sums in the following two aspects:

- the modulus q is now an arbitrary positive integer;
- the weights are supported on an arbitrary subset of $\mathcal{M} \subseteq \mathbb{Z}_q^*$.

Using the approach of [24], augmented with several new arguments, we improve and generalise previous bounds on these sums. It is important to note, that our method does not rely on algebraic geometry results, as the method of [1, 8, 18] and thus work modulo composite numbers as well as modulo primes.

Furthermore, we use similar ideas to the weighted sums of Gauss sums

$$\mathcal{G}_q(\chi, n) = \sum_{x \in \mathbb{Z}_q^*} \chi(x) \mathbf{e}_q(nx),$$

where χ is a Dirichlet character modulo q , see [14, Chapter 3] for a background on characters. More precisely, given a sequence of weights $\mathcal{W} = \{\omega_\chi\}_{\chi \in \Xi}$, supported on a subset Ξ of the set Ω_q^* of primitive Dirichlet character modulo q . we define the weighted sums of Gauss sums

$$\mathcal{T}_q(\mathcal{W}; \Xi, \mathcal{J}) = \sum_{\chi \in \Xi} \sum_{n \in \mathcal{J}} \omega_\chi \mathcal{G}_q(\chi, n).$$

where, as before, $\mathcal{J} = \{L+1, \dots, L+N\} \subseteq [1, q-1]$. We certainly have the full analogue of (1.1):

$$|\mathcal{T}_q(\mathcal{W}; \Xi, \mathcal{J})| \leq Nq^{1/2+o(1)} \sum_{\chi \in \Xi} |\omega_\chi|,$$

see [14, Equation (3.14)]. The study of cancellations between Gauss sums has been initiated by Katz and Zheng [16], in a different form as a result about the uniformity of distribution of their arguments. Furthermore, in the case of prime $q = p$ and constant weights, one can obtain a nontrivial upper bound on

$$\mathcal{T}_p(\{1\}_{\chi \in \Xi}; \Xi, \mathcal{J}) = \sum_{\chi \in \Xi} \sum_{n \in \mathcal{J}} \mathcal{G}_p(\chi, n)$$

under the condition $MN \geq p^{1+\varepsilon}$ for a fixed $\varepsilon > 0$, where $M = \#\Xi$, from the uniformity of distribution result of [23, Theorem 3]. Here we obtain stronger and more general bounds.

As an application of our new bounds with Kloosterman sums, we also improve the power saving in the error term of the asymptotic formula for mixed moments of L -series associated with Hecke eigenforms, which refines the previous result of Blomer, Fouvry, Kowalski, Michel and Milićević [1, Theorem 1.2].

1.2. General notation. We define the norms

$$\|\mathcal{A}\|_\infty = \max_{m \in \mathcal{M}} |\alpha_m| \quad \text{and} \quad \|\mathcal{A}\|_\sigma = \left(\sum_{m \in \mathcal{M}} |\alpha_m|^\sigma \right)^{1/\sigma},$$

where $\sigma > 0$.

We always assume that the sequence of weights $\mathcal{A} = \{\alpha_m\}_{m \in \mathcal{I}}$ is supported only on m with $\gcd(m, q) = 1$, that is, we have $\alpha_m = 0$ if $\gcd(m, q) > 1$.

Throughout the paper, as usual $A \ll B$ is equivalent to the inequality $|A| \leq cB$ with some constant $c > 0$, which occasionally, where obvious, may depend on the real parameter $\varepsilon > 0$ and on the integer parameter $r \geq 1$, and is absolute otherwise.

The letter p always denotes a prime number.

1.3. Previous results. For

$$\mathcal{I} = \{K + 1, \dots, K + M\}, \quad \mathcal{J} = \{L + 1, \dots, L + N\} \subseteq [1, q - 1],$$

the sums $\mathcal{S}_p(\mathcal{A}; \mathcal{I}, \mathcal{J})$ have been estimated by Fouvry, Kowalski and Michel [8, Theorem 1.17] as a part of a much more general result about sums of so-called *trace functions*. For example, by [8, Theorem 1.17(2)], for initial intervals $\mathcal{I} = \{1, \dots, M\}$ and $\mathcal{J} = \{1, \dots, N\}$, we have

$$(1.2) \quad |\mathcal{S}_p(\mathcal{A}; \mathcal{I}, \mathcal{J})| \leq \|\mathcal{A}\|_1 p^{1+o(1)}.$$

Furthermore, by a result of Blomer, Fouvry, Kowalski, Michel and Milićević [1, Theorem 6.1], also for an initial interval \mathcal{I} and an arbitrary interval \mathcal{J} with

$$(1.3) \quad MN \leq p^{3/2} \quad \text{and} \quad M \leq N^2,$$

we have

$$(1.4) \quad |\mathcal{S}_p(\mathcal{A}; \mathcal{I}, \mathcal{J})| \leq (\|\mathcal{A}\|_1 \|\mathcal{A}\|_2)^{1/2} M^{1/12} N^{7/12} p^{3/4+o(1)}.$$

The results of [1, 8] are based on deep methods originating from algebraic geometry, such the *Weil and Deligne bounds*, see [14, Chapter 11]. A much more elementary approach, suggested in [24], yields the estimate

$$(1.5) \quad \mathcal{S}_p(\mathcal{A}; \mathcal{I}, \mathcal{J}) \ll \|\mathcal{A}\|_2 N^{1/2} p.$$

In particular, we see that the approach of [24] improves the bounds from [1, 8] for

$$N < Mp^{-\varepsilon} \quad \text{and} \quad M^4 N \geq p^{3+\varepsilon}$$

with any fixed $\varepsilon > 0$.

For the purpose of comparison between previous results and our new bounds, we rewriting the bounds (1.1), (1.2), (1.4) and (1.5) in terms of $\|\mathcal{A}\|_\infty$ and combine them in one bound

$$(1.6) \quad \mathcal{S}_p(\mathcal{A}; \mathcal{I}, \mathcal{J}) \ll \|\mathcal{A}\|_\infty \min\{MNp^{1/2}, Mp, M^{5/6}N^{7/12}p^{3/4}, M^{1/2}N^{1/2}p\}p^{o(1)},$$

where we also ignore the necessary condition (1.3) for the bound (1.4) to apply.

For sums of Gauss sums, no general results have been known. However, for a prime $q = p$ as we have mentioned, one can derive a non-trivial bound in the case of constant weights from [23, Theorem 3] and in fact for more general sums with the summation over n over an arbitrary set $\mathcal{N} \subseteq \mathbb{Z}_p$.

2. NEW RESULTS

2.1. Bounds of sums of Kloosterman sum. We remark that our bounds only involve the norms of the weights \mathcal{A} but do not explicitly depend on the size of the set \mathcal{M} on which they are supported. Hence, without loss of generality, we can assume that $\mathcal{M} = \mathbb{Z}_q^*$ and thus we simplify the notation as

$$\mathcal{S}_q(\mathcal{A}; \mathcal{J}) = \sum_{m \in \mathbb{Z}_q^*} \sum_{n \in \mathcal{J}} \alpha_m \mathcal{K}_q(m, n).$$

Theorem 2.1. *For any integer $q \geq 1$, we have,*

$$\mathcal{S}_q(\mathcal{A}; \mathcal{J}) \ll (\|\mathcal{A}\|_1 \|\mathcal{A}\|_2)^{1/2} (N^{1/8} q + N^{1/2} q^{3/4}) q^{o(1)}.$$

Returning to our original settings and assuming that the sequence of weights \mathcal{A} is supported on a set $\mathcal{M} \subseteq \mathbb{Z}_q^*$ of size M , we can rewrite that bound of Theorem 2.1 in terms of $\|\mathcal{A}\|_\infty$ as

$$(2.1) \quad \mathcal{S}_q(\mathcal{A}; \mathcal{M}, \mathcal{J}) \ll \|\mathcal{A}\|_\infty M^{3/4} (N^{1/8} q + N^{1/2} q^{3/4}) q^{o(1)}.$$

We see that the bound (2.1), besides being more general than (1.6), also gives a better result, provided that

$$(2.2) \quad M^2 N^7 \geq q^4, \quad M^2 N^{11} \geq q^6, \quad qM \geq N^2, \quad N^3 \geq M^2 \geq N.$$

Writing $M = q^\mu$ and $N = q^\nu$ we see that the conditions (2.2) define a polygon with vertices

$$(1/4, 1/2), \quad (1/3, 2/3), \quad (1, 1), \quad (1, 2/3), \quad (9/14, 3/7),$$

in the (μ, ν) -plane, see also Figure 1. The most important for applications fact is that the point $(1/2, 1/2)$ is an interior point of this polygon. On the other hand, we also remark that although we have ignored the restriction (1.3) for the bound (1.4) to hold, taking it into account does not increase the region where (2.1) improves the previously known bounds.

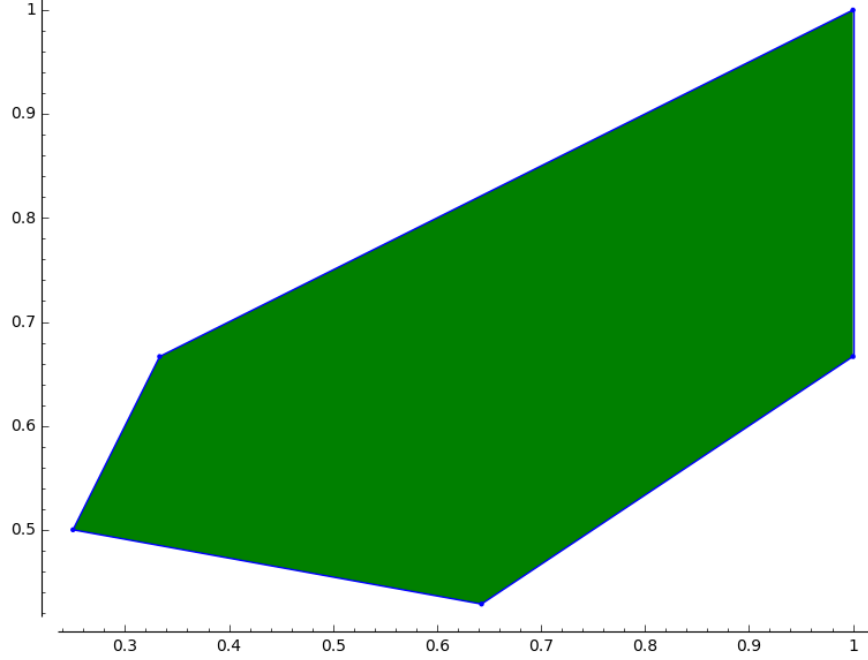


FIGURE 1. Polygon of relative sizes of M and N defined by (2.2).

In particular, even ignoring the differences in the generality, in the situation where all previous bounds apply, when $M \sim N \sim p^{1/2}$ (which is a critical range for several applications), Theorem 2.1 implies the bound $p^{3/2-1/16+o(1)}$ instead of $p^{3/2-1/24+o(1)}$ given by (1.6) (which actually comes from (1.4)).

We also obtain a stronger bound on average over q in a dyadic interval $[Q, 2Q]$:

Theorem 2.2. *For any fixed real $\varepsilon > 0$ and integer $r \geq 2$, for any sufficiently large $Q \geq 1$, for all but at most $Q^{1-2r\varepsilon+o(1)}$ integers $q \in [Q, 2Q]$ we have*

$$\mathcal{S}_q(\mathcal{A}; \mathcal{J}) \ll \|\mathcal{A}\|_1^{1-1/r} \|\mathcal{A}\|_2^{1/r} (q + N^{1/2} q^{1/2+1/2r}) q^{\varepsilon+o(1)}.$$

Again, when the sequence of weights \mathcal{A} is supported on a set $\mathcal{M} \subseteq \mathbb{Z}_q^*$ of size M , we can rewrite that bound of Theorem 2.2 in terms of $\|\mathcal{A}\|_\infty$ as

$$\mathcal{S}_q(\mathcal{A}; \mathcal{M}, \mathcal{J}) \ll \|\mathcal{A}\|_\infty M^{1-1/2r} (q + N^{1/2} q^{1/2+1/2r}) q^{\varepsilon+o(1)}.$$

2.2. Bounds of sums of Gauss sum. For the sums of Gauss sums we also have bounds that only involve the norms of the weights \mathcal{W} but do not explicitly depend on the size of the set Ξ on which they are

supported. Hence, without loss of generality, we can assume that $\Xi = \Omega_q^*$ and thus we simplify the notation as

$$\mathcal{T}_q(\mathcal{W}; \mathcal{J}) = \sum_{\chi \in \Omega_q^*} \sum_{n \in \mathcal{J}} \omega_\chi \mathcal{G}_q(\chi, n).$$

Theorem 2.3. *For any integer $q \geq 1$, we have,*

$$\mathcal{T}_q(\mathcal{W}; \mathcal{J}) \ll (\|\mathcal{W}\|_1 \|\mathcal{W}\|_2)^{1/2} (q + N^{1/2} q^{3/4}) q^{o(1)}.$$

Finally, we also have

Theorem 2.4. *For any fixed real $\varepsilon > 0$ and integer $r \geq 2$, for any sufficiently large $Q \geq 1$, for all but at most $Q^{1-2r\varepsilon+o(1)}$ integers $q \in [Q, 2Q]$ we have*

$$\mathcal{T}_q(\mathcal{W}; \mathcal{J}) \ll \|\mathcal{W}\|_1^{1-1/r} \|\mathcal{W}\|_2^{1/r} (q + N^{1/2} q^{1/2+1/2r}) q^{\varepsilon+o(1)}.$$

If the sequence of weights \mathcal{W} is supported on a set $\Xi \subseteq \Omega_q^*$ of size M , then we have

$$\mathcal{T}_q(\mathcal{W}; \Xi, \mathcal{J}) \ll \|\mathcal{W}\|_\infty M^{1-1/2r} (q + N^{1/2} q^{1/2+1/2r}) q^{\varepsilon+o(1)}.$$

where we can take $r = 2$ under the conditions of Theorem 2.3 and any integer $r \geq 2$ under the conditions of Theorem 2.4.

2.3. Applications. Inserting the bound of Theorem 2.1 (in the special case of a prime $q = p$ and the weights supported on the initial interval $\mathcal{I}_0 = \{1, \dots, M\}$) in the argument of the proof of [1, Theorem 1.2], we improve the error term in the asymptotic formula for mixed moments of L -series associated with Hecke eigenforms.

To be more precise we need to recall some definitions from [1].

Let f be a Hecke eigenform and let χ be a primitive Dirichlet character modulo q . We define the following L -series by the absolutely converging for $\Re s > 1$ series:

$$L(f \otimes \chi, s) = \sum_{n=1}^{\infty} \frac{\lambda_f(n) \chi(n)}{n^s} \quad \text{and} \quad L(f, s) = \sum_{n=1}^{\infty} \frac{\lambda_f(n)}{n^s}, \quad \Re s > 1,$$

where $\{\lambda_f(n)\}_{n=1}^{\infty}$ are the Hecke eigenvalues of f . We also define the partial derivative at $s = 1/2$

$$F(z) = \left. \frac{\partial}{\partial s} \right|_{s=1/2} E(z; s)$$

of the *Eisenstein series* defined by

$$E(z; s) = \frac{1}{2} \sum_{\substack{c, d \in \mathbb{Z} \\ \gcd(c, d)=1}} \frac{(\Im z)^s}{|cz + d|^{2s}}, \quad \Re s > 1, \quad \Im z > 0,$$

Finally, we define the average value

$$\mathfrak{M}_{f,F}(q) = \frac{1}{\#\Omega_q^*} \sum_{\chi \in \Omega_q^*} L(f \otimes \chi, 1/2) \overline{L(F \otimes \chi, 1/2)},$$

where, as before, Ω_q^* denotes the set of all primitive Dirichlet characters modulo q . Clearly, $\#\Omega_p^* = p - 2$ for a prime p . As usual, $\zeta(s)$ denotes the Riemann zeta-function.

Theorem 2.5. *For any prime p , we have,*

$$\mathfrak{M}_{f,F}(p) = \frac{L(f, 1)^2}{\zeta(2)} + O\left(p^{-1/64+o(1)}\right),$$

where the implied constant depends only on f .

Theorem 2.5 improves the error term $p^{-1/68+o(1)}$ of [1, Theorem 1.2].

3. PRELIMINARIES

3.1. Equations and congruences with reciprocals. An important tool in our argument is an upper bound on the number of solutions $J_r(q; K)$ to the congruence

$$\frac{1}{x_1} + \dots + \frac{1}{x_r} \equiv \frac{1}{x_{r+1}} + \dots + \frac{1}{x_{2r}} \pmod{q}, \quad 1 \leq x_1, \dots, x_{2r} \leq K.$$

where $r = 1, 2, \dots$.

For arbitrary q and K , good upper bounds on $J_r(q; K)$ are known only for $r = 2$ (and of course in the trivial case $r = 1$) and are due to Heath-Brown [12, Page 368] (see the bound on the sums of quantities $m(s)^2$ in the notation of [12]). More precisely, we have:

Lemma 3.1. *For $1 \leq K \leq q$ we have*

$$J_2(q; K) \leq (K^{7/2}q^{-1/2} + K^2) q^{o(1)}.$$

It is also shown by Fouvry and Shparlinski [10, Lemma 2.3] that the bound of Lemma 3.1 can be improved on average over q in a dyadic interval $[Q, 2Q]$. The same argument also works for $J_r(q; K)$ without any changes.

Indeed, let $J_r(K)$ be the number of solutions to the equation

$$\frac{1}{x_1} + \dots + \frac{1}{x_r} = \frac{1}{x_{r+1}} + \dots + \frac{1}{x_{2r}}, \quad 1 \leq x_1, \dots, x_{2r} \leq K,$$

where $r = 1, 2, \dots$. We recall that by the result of Karatsuba [15] (presented in the proof of [15, Theorem 1]), see also [3, Lemma 4], we have:

Lemma 3.2. *For any fixed positive integer r , we have*

$$J_r(K) \leq K^{r+o(1)}.$$

Now repeating the argument of the proof of [10, Lemma 2.3] and using Lemma 3.2 in the appropriate place, we obtain:

Lemma 3.3. *For any fixed positive integer r and sufficiently large integers $1 \leq K \leq Q$, we have*

$$\frac{1}{Q} \sum_{Q \leq q \leq 2Q} J_{r,q}(K) \leq (K^{2r} Q^{-1} + K^r) Q^{o(1)}.$$

3.2. Equations and congruences with products. Let $R_r(q; K)$ be the number of solutions to the congruence

$$x_1 \dots x_r \equiv x_{r+1} \dots x_{2r} \pmod{q}, \quad 1 \leq x_1, \dots, x_{2r} \leq K.$$

where $r = 1, 2, \dots$. The proof of Theorem 2.3 uses the bound of Friedlander and Iwaniec [11, Lemma 3] on $R_2(q; K)$ (which is formulated as a bound on the 4th moment of character sums), see also [7, Theorem 2]. We present it in a simplified form.

Lemma 3.4. *For $1 \leq K \leq q$ we have*

$$R_2(q; K) \leq (K^4 q^{-1} + K^2) q^{o(1)}.$$

Now let $R_r(K)$ be the number of solutions to the equation

$$x_1 \dots x_r = x_{r+1} \dots x_{2r}, \quad 1 \leq x_1, \dots, x_{2r} \leq K,$$

where $r = 1, 2, \dots$. We classical bound on the divisor function immediately implies:

Lemma 3.5. *For any fixed positive integer r , we have*

$$R_r(K) \leq K^{r+o(1)}.$$

Thus using Lemma 3.5 instead of Lemma 3.2 we obtain an analogue of Lemma 3.3:

Lemma 3.6. *For any fixed positive integer r and sufficiently large integers $1 \leq K \leq Q$, we have*

$$\frac{1}{Q} \sum_{Q \leq q \leq 2Q} R_{r,q}(K) \leq (K^{2r} Q^{-1} + K^r) Q^{o(1)}.$$

4. PROOFS

4.1. **Proof of Theorem 2.1.** For an integer u we define

$$\langle u \rangle_q = \min_{k \in \mathbb{Z}} |u - kq|$$

as the distance to the closest integer, which is a multiple of q .

Changing the order of summation and then changing the variable $x \mapsto \bar{x}$, we obtain

$$\begin{aligned} \mathcal{S}_q(\mathcal{A}; \mathcal{J}) &= \sum_{x=1}^{p-1} \sum_{m \in \mathbb{Z}_q^*} \alpha_m \mathbf{e}_q(mx) \sum_{n \in \mathcal{J}} \mathbf{e}_q(n\bar{x}) \\ &= \sum_{x \in \mathbb{Z}_q^*} \sum_{m \in \mathbb{Z}_q^*} \alpha_m \mathbf{e}_q(m\bar{x}) \sum_{n \in \mathcal{J}} \mathbf{e}_q(nx). \end{aligned}$$

Hence

$$\mathcal{S}_q(\mathcal{A}; \mathcal{J}) = \sum_{m \in \mathbb{Z}_q^*} \sum_{x \in \mathbb{Z}_q^*} \alpha_m \gamma_x \mathbf{e}_q(m\bar{x}),$$

where

$$|\gamma_x| \leq \min \left\{ N, \frac{q}{\langle u \rangle_q} \right\}.$$

We now set $I = \lceil \log(N/2) \rceil$ and define $2(I+1)$ the sets

$$\mathcal{X}_0^\pm = \{x \in \mathbb{Z} : 0 < \pm x \leq q/N\},$$

$$\mathcal{X}_i^\pm = \{x \in \mathbb{Z} : \min\{q/2, e^i q/N\} \geq \pm x > e^{i-1} q/N\}, \quad i = 1, \dots, I.$$

Therefore,

$$(4.1) \quad \mathcal{S}_q(\mathcal{A}; \mathcal{J}) \ll \sum_{i=0}^I (|S_i^+| + |S_i^-|),$$

where

$$S_i^\pm = \sum_{m \in \mathbb{Z}_q^*} \sum_{x \in \mathcal{X}_i^\pm} \alpha_m \gamma_x \mathbf{e}_q(m\bar{x}), \quad i = 0, \dots, I.$$

Now we present the argument in a general form with an arbitrary $r \geq 2$. We then apply it with $r = 2$ since we use Lemma 3.1. However in the proof of Theorem 2.2 we use it in full generality.

Let us fix some integer $r \geq 2$. Writing

$$|S_i^\pm| \leq \sum_{m \in \mathbb{Z}_q^*} |\alpha_m|^{(r-1)/r} |\alpha_m^2|^{1/2r} \left| \sum_{x \in \mathcal{X}_i^\pm} \alpha_m \gamma_x \mathbf{e}_q(m\bar{x}) \right|,$$

by the Hölder inequality, for every $i = 0, \dots, I$ and every choice of the sign ‘+’ or ‘−’, we obtain

$$\begin{aligned}
 |S_i^\pm| &\leq \left(\sum_{m \in \mathbb{Z}_q^*} |\alpha_m| \right)^{1-1/r} \left(\sum_{m \in \mathbb{Z}_q^*} |\alpha_m|^2 \right)^{1/2r} \\
 (4.2) \quad &\left(\sum_{m \in \mathbb{Z}_q^*} \left| \sum_{x \in \mathcal{X}_i^\pm} \gamma_x \mathbf{e}_q(m\bar{x}) \right|^{2r} \right)^{1/2r} \\
 &= \|\mathcal{A}\|_1^{1-1/r} \|\mathcal{A}\|_2^{1/r} \left(\sum_{m \in \mathbb{Z}_q^*} \left| \sum_{x \in \mathcal{X}_i^\pm} \gamma_x \mathbf{e}_q(m\bar{x}) \right|^{2r} \right)^{1/2r}.
 \end{aligned}$$

Extending the summation over m to the whole ring \mathbb{Z}_q , opening up the inner sum, changing the order of summation and using the orthogonality of exponential functions, we obtain

$$\begin{aligned}
 &\sum_{m \in \mathbb{Z}_q^*} \left| \sum_{x \in \mathcal{X}_i^\pm} \gamma_x \mathbf{e}_q(m\bar{x}) \right|^{2r} \\
 &\leq \sum_{m \in \mathbb{Z}_q} \sum_{x_1, \dots, x_{2r} \in \mathcal{X}_i^\pm} \prod_{j=1}^r \gamma_{x_j} \overline{\gamma_{x_{r+j}}} \mathbf{e}_q \left(m \sum_{j=1}^r (\bar{x}_j - \bar{x}_{r+j}) \right) \\
 &\leq \sum_{x_1, \dots, x_{2r} \in \mathcal{X}_i^\pm} \prod_{j=1}^r \gamma_{x_j} \overline{\gamma_{x_{r+j}}} \sum_{m \in \mathbb{Z}_q} \mathbf{e}_q \left(m \sum_{j=1}^r (\bar{x}_j - \bar{x}_{r+j}) \right) \\
 &= q \sum_{\substack{x_1, \dots, x_{2r} \in \mathcal{X}_i^\pm \\ \bar{x}_1 + \dots + \bar{x}_r \equiv \bar{x}_{r+1} + \dots + \bar{x}_{2r} \pmod{q}}} \prod_{j=1}^r \gamma_{x_j} \overline{\gamma_{x_{r+j}}}.
 \end{aligned}$$

We also observe that for $x \in \mathcal{X}_i^\pm$ we have

$$|\gamma_x| \ll e^{-i} N.$$

Hence, for every $i = 0, \dots, I$ we have

$$\begin{aligned}
 (4.3) \quad & \sum_{m \in \mathbb{Z}_q^*} \left| \sum_{x \in \mathcal{X}_i^\pm} \gamma_x \mathbf{e}_q(m\bar{x}) \right|^{2r} \\
 & \ll e^{-2ri} N^{2r} q \sum_{\substack{x_1, \dots, x_{2r} \in \mathcal{X}_i^\pm \\ \bar{x}_1 + \dots + \bar{x}_r \equiv \bar{x}_{r+1} + \dots + \bar{x}_{2r} \pmod{q}}} 1 \\
 & \leq e^{-2ri} N^{2r} q J_r(q; \lfloor e^i q / N \rfloor).
 \end{aligned}$$

Now using (4.3) with $r = 2$ and revoking Lemma 3.1 we obtain

$$\sum_{m \in \mathbb{Z}_q^*} \left| \sum_{x \in \mathcal{X}_0^\pm} \gamma_x \mathbf{e}_q(m\bar{x}) \right|^4 \leq e^{-4i} N^4 (N^{-7/2} q^3 + N^{-2} q^2) q,$$

Nw we see from (4.2) that

$$\begin{aligned}
 (4.4) \quad & |S_i^\pm| \leq (\|\mathcal{A}\|_1 \|\mathcal{A}\|_2)^{1/2} \\
 & e^{-i} N (e^{7i/2} N^{-7/2} q^3 + e^{2i} N^{-2} q^2)^{1/4} q^{1/4+o(1)} \\
 & \leq e^{-i/8} (\|\mathcal{A}\|_1 \|\mathcal{A}\|_2)^{1/2} (N^{1/8} q + N^{1/2} q^{3/4}) q^{o(1)}.
 \end{aligned}$$

Therefore,

$$(4.5) \quad \sum_{i=0}^I |S_i^\pm| \leq (\|\mathcal{A}\|_1 \|\mathcal{A}\|_2)^{1/2} (N^{1/8} q + N^{1/2} q^{3/4}) q^{o(1)}.$$

Substituting (4.5) in (4.1), we obtain the result.

4.2. Proof of Theorem 2.2. We proceed as in the proof of Theorem 2.1, in particular, we set $I = \lceil \log(N/2) \rceil$. We also define $K_i = \lfloor 2e^i Q / N \rfloor$ and replace $J_r(q; \lfloor e^i q / N \rfloor)$ with $J_r(q; K_i)$ in (4.3), $i = 0, \dots, I$.

We know see that by Lemma 3.3 for every $i = 0, \dots, I$ for all but at most $Q^{1-2r\epsilon+o(1)}$ integers $q \in [Q, 2Q]$ we have

$$(4.6) \quad J_{r,q}(K_i) \leq (K_i^{2r} q^{-1} + K_i^r) Q^{2r\epsilon}.$$

Since $I = Q^{o(1)}$, we see that for all but at most $Q^{1-2r\epsilon+o(1)}$ integers $q \in [Q, 2Q]$, the bound (4.6) holds for all $i = 0, \dots, I$ simultaneously. For every such q , using (4.6), instead of the bound of Lemma 3.1, we obtain

$$\begin{aligned}
 |S_i^\pm| & \ll \|\mathcal{A}\|_1^{1-1/r} \|\mathcal{A}\|_2^{1/r} e^{-i} N (e^{2ri} N^{-2r} q^{2r-1} + e^{ri} N^{-r} q^r)^{1/2r} q^{1/2r} Q^\epsilon \\
 & \ll \|\mathcal{A}\|_1^{1-1/r} \|\mathcal{A}\|_2^{1/r} (q + N^{1/2} q^{1/2+1/2r}) Q^\epsilon
 \end{aligned}$$

instead of (4.4) for every $i = 0, \dots, I$. Since $I = Q^{o(1)}$, the result now follows.

4.3. Proof of Theorem 2.3. We define I , the sets \mathcal{X}_i^\pm , $i = 0, \dots, I$, and the quantities γ_x , $x \in \mathbb{Z}_q^*$, as in the proof of Theorem 2.1. Therefore, instead of (4.1) we have

$$(4.7) \quad \mathcal{T}_q(\mathcal{W}; \mathcal{J}) \ll \sum_{i=0}^I (|T_i^+| + |T_i^-|),$$

where

$$T_i^\pm = \sum_{\chi \in \Omega_q^*} \sum_{x \in \mathcal{X}_i^\pm} \omega_\chi \gamma_x \chi(x), \quad i = 0, \dots, I.$$

Furthermore, we have the following analogue of (4.2)

$$(4.8) \quad |T_i^\pm| \leq \|\mathcal{W}\|_1^{1-1/r} \|\mathcal{W}\|_2^{1/r} \left(\sum_{\chi \in \Omega_q^*} \left| \sum_{x \in \mathcal{X}_i^\pm} \gamma_x \chi(x) \right|^{2r} \right)^{1/2r},$$

from which, by the orthogonality of characters, we derive an analogue of (4.3). More precisely, for every $i = 0, \dots, I$ we obtain

$$(4.9) \quad \sum_{\chi \in \Omega_q^*} \left| \sum_{x \in \mathcal{X}_i^\pm} \gamma_x \chi(x) \right|^{2r} \leq e^{-2ri} N^{2r} q R_r(q; \lfloor e^i q / N \rfloor).$$

Now using (4.9) with $r = 2$ and revoking Lemma 3.4 we obtain

$$\sum_{m \in \mathbb{Z}_q^*} \left| \sum_{x \in \mathcal{X}_0^\pm} \gamma_x \mathbf{e}_q(m\bar{x}) \right|^4 \leq e^{-4i} N^4 (N^{-4} q^3 + N^{-2} q^2) q,$$

and thus by (4.8), we have

$$\begin{aligned} |T_i^\pm| &\leq (\|\mathcal{W}\|_1 \|\mathcal{W}\|_2)^{1/2} \\ &\quad e^{-i} N (e^{4i} N^{-4} q^3 + e^{2i} N^{-2} q^2)^{1/4} q^{1/4+o(1)} \\ &\leq (\|\mathcal{W}\|_1 \|\mathcal{W}\|_2)^{1/2} (q + N^{1/2} q^{3/4}) q^{o(1)}. \end{aligned}$$

Therefore,

$$(4.10) \quad \sum_{i=0}^I |T_i^\pm| \leq (\|\mathcal{W}\|_1 \|\mathcal{W}\|_2)^{1/2} (q + N^{1/2} q^{3/4}) q^{o(1)}.$$

Substituting (4.10) in (4.7), we obtain the result.

4.4. Proof of Theorem 2.4. We proceed as in the proof of Theorem 2.3, using Lemma 3.6 instead of Lemma 3.4 in the appropriate place (see also the proof of Theorem 2.2).

4.5. Proof of Theorem 2.5. We simply incorporate the bound of Theorem 2.1 in the arsenal of bounds used in [1, Section 6.4.1]. Augmenting the `Mathematica` code, provided in [1, Section 7.4.1], with this new bound, we see that the contribution from the terms considered in [1, Section 6.4.1] can be estimated as $p^{-1/52+o(1)}$. Hence, now the error term is dominated by the terms treated in [1, Section 6.4.2], which contribute at most $p^{-1/64+o(1)}$. The result now follows.

5. COMMENTS AND FURTHER APPLICATIONS

We note that the error term of Theorem 2.5 is now dominated by the terms whose treatment is free of any use of sums of Kloosterman sums, see [1, Section 6.4.2]. Hence no further improvement is possible until this part is refined. For example, at the moment we cannot take any advantage of Theorem 2.2 in this context.

We also recall that for a prime $q = p$ and small K , a series of bounds on $J_r(p, K)$ have been given by Bourgain and Garaev [3, 4]. These bounds can also be used in the argument of the proof of Theorem 2.1, leading to a series of estimates when N is close to q . Similarly, the bounds of [5, 6] on $R_r(p, K)$ can be used in the argument of the proof of Theorem 2.3.

We note the suggested here approach can be applied to many other families of bilinear sums of the form

$$\mathcal{S}_{k,q}(\mathcal{A}; \mathcal{J}) = \sum_{m \in \mathbb{Z}_q^*} \sum_{n \in \mathcal{J}} \alpha_m \sum_{x \in \mathbb{Z}_q^*} \mathbf{e}_q(mx^{-k} + nx)$$

for an integer $k \geq 1$ (generalising the sums $\mathcal{S}_{1,q}(\mathcal{A}; \mathcal{J}) = \mathcal{S}_q(\mathcal{A}; \mathcal{J})$). Indeed, instead of Lemma 3.1, in the appropriate place of the argument, one simply uses the bound of Heath-Brown [13, Lemma 1] for $k = 2$ and a more general bound of Pierce [21, Theorem 4] for arbitrary integer $k \geq 1$, see also [3, Proposition 1]. We note that the sums $\mathcal{S}_{2,q}(\mathcal{A}; \mathcal{J})$ and related sums have been estimated by Nunes [20, Theorems 1.2 and 1.3] as a tool in investigating the distribution of squarefree numbers in arithmetic progressions. Nunes [20] uses the method of [1, 18] and it is very plausible that the method of this work, and also the method of [24] in the case of constant weights, may lead to stronger results.

Finally, it is easy to see that one can also use the same approach to estimate the sums

$$\sum_{m \in \mathbb{Z}_p^*} \sum_{n \in \mathcal{J}} \alpha_m \sum_{x \in \mathbb{Z}_t} \mathbf{e}_p(mg^x) \mathbf{e}_t(nx),$$

where g is an integer of multiplicative order t modulo a prime p and $\mathcal{J} = \{L + 1, \dots, L + N\} \subseteq [1, t]$. In this case, instead of Lemma 3.1, one uses the bound

$$\#\{1 \leq u, v, x, y \leq K : g^u + g^v \equiv g^x + g^y \pmod{p}\} \ll K^{5/2},$$

for any $K \leq t$, which follows immediately from a result of Roche-Newton, Rudnev and Shkredov [22, Theorem 6] (see also the proof of [22, Theorem 18]).

ACKNOWLEDGEMENT

The authors would like to thank Ramon Nunes for useful discussions and in particular for the information about the results of Heath-Brown [13, Lemma 1] and Pierce [21, Theorem 4].

This work was supported by ARC Grant DP140100118.

REFERENCES

- [1] V. Blomer, É. Fouvry, E. Kowalski, P. Michel and D. Milićević, ‘On moments of twisted L -functions’, *Amer. J. of Math.*, (to appear).
- [2] V. Blomer, É. Fouvry, E. Kowalski, P. Michel and D. Milićević, ‘Some applications of smooth bilinear forms with Kloosterman sums’, *Proc. Steklov Math. Inst.*, (to appear).
- [3] J. Bourgain and M. Z. Garaev, ‘Sumsets of reciprocals in prime fields and multilinear Kloosterman sums’, *Izv. Ross. Akad. Nauk Ser. Mat.*, **78** (2014), 9–72 (in Russian); translation in *Izv. Math.*, **78** (2014), 656–707.
- [4] J. Bourgain and M. Z. Garaev, ‘Kloosterman sums in residue rings’, *Acta Arith.*, **164** (2014), 43–64.
- [5] J. Bourgain, M. Z. Garaev, S. V. Konyagin and I. E. Shparlinski, ‘On congruences with products of variables from short intervals and applications’, *Proc. Steklov Math. Inst.*, **280** (2013), 67–96.
- [6] J. Bourgain, M. Z. Garaev, S. V. Konyagin and I. E. Shparlinski, ‘Multiplicative congruences with variables from short intervals’, *J. d’Analyse Math.*, **124** (2014), 117–147.
- [7] T. Cochrane and S. Shi, ‘The congruence $x_1x_2 \equiv x_3x_4 \pmod{m}$ and mean values of character sums’, *J. Number Theory*, **130** (2010), 767–785.
- [8] É. Fouvry, E. Kowalski and P. Michel, ‘Algebraic trace functions over the primes’, *Duke Math. J.*, **163** (2014), 1683–1736.
- [9] É. Fouvry, E. Kowalski, P. Michel, C. S. Raju, J. Rivat and K. Soundararajan, ‘On short sums of trace functions’, *Preprint*, 2015 (available from <http://arxiv.org/abs/1508.00512>).

- [10] É. Fouvry and I.E. Shparlinski, ‘On a ternary quadratic form over primes’, *Acta Arith.*, **150** (2011), 285–314.
- [11] J. B. Friedlander and H. Iwaniec, ‘The divisor problem for arithmetic progressions’, *Acta Arith.*, **45** (1985), 273–277.
- [12] D.R. Heath-Brown, ‘Almost primes in arithmetic progressions and short intervals’, *Math. Proc. Cambridge Philos. Soc.*, **83** (1978), 357–375.
- [13] D.R. Heath-Brown, ‘The least square-free number in an arithmetic progression’, *J. Reine Angew. Math.*, **332** (1982), 204–220.
- [14] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society, Providence, RI, 2004.
- [15] A. A. Karatsuba, ‘Analogues of Kloosterman sums’, *Izv. Ross. Akad. Nauk Ser. Mat. (Transl. as Russian Acad. Sci. Izv. Math.)*, **55**(5) (1995), 93–102 (in Russian)
- [16] N. M. Katz and Z. Zheng, ‘On the uniform distribution of Gauss sums and Jacobi sums’, *Analytic Number Theory, Allerton Park 1995*, Progress in Mathematics, vol. 139, Birkhäuser, Basel, 1996, 537–558.
- [17] R. Khan, ‘The divisor function in arithmetic progressions modulo prime powers’, *Mathematika*, **62** (2016), 898–908.
- [18] E. Kowalski, P. Michel and W. Sawin, ‘Bilinear forms with Kloosterman sums and applications’, *Preprint*, 2015 (available from <http://arxiv.org/abs/1511.01636>).
- [19] K. Liu, I. E. Shparlinski and T. P. Zhang, ‘Divisor problem in arithmetic progressions modulo a prime power’, *Preprint*, 2016 (available from <http://arxiv.org/abs/1602.03583>).
- [20] R. M. Nunes, ‘Squarefree numbers in large arithmetic progressions’, *Preprint*, 2016 (available from <http://arxiv.org/abs/1602.00311>).
- [21] L. B. Pierce, ‘The 3-part of class numbers of quadratic fields’, *J. London Math. Soc.*, **71** (2005), 579–598.
- [22] O. Roche-Newton, M. Rudnev and I. D. Shkredov, ‘New sum-product type estimates over finite fields’, *Adv. Math.*, **293** (2016), 589–605.
- [23] I. E. Shparlinski, ‘On the distribution of arguments of Gauss sums’, *Kodai Math. J.*, **32** (2009), 172–177.
- [24] I. E. Shparlinski and T. P. Zhang, ‘Cancellations amongst Kloosterman sums’, *Acta Arith.*, (to appear).
- [25] J. Wu and P. Xi, ‘Arithmetic exponent pairs for algebraic trace functions and applications’, *Preprint*, 2016 (available from <http://arxiv.org/abs/1603.07060>).
- [26] P. Xi (with an appendix by É. Fouvry, E. Kowalski and P. Michel) ‘Large sieve inequalities for algebraic trace functions’, *Intern. Math. Res. Notices*, (to appear).

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF NEW SOUTH WALES,
 SYDNEY, NSW 2052, AUSTRALIA

E-mail address: igor.shparlinski@unsw.edu.au